Management and Production Engineering Review
Volume 11 • Number 3 • September 2020 • pp. 38–47
DOI: 10.24425/mper.2020.134930

BIG DATA-DRIVEN FRAMEWORK FOR VIRAL CHURN PREVENTION: A CASE STUDY

Laura Lucantoni, Sara Antomarioni, Maurizio Bevilacqua, Filippo Emanuele Ciarapica

Dipartimento di Ingegneria Industriale e Scienze Matematiche, Università Politecnica Delle Marche, Italy

Corresponding author:
Maurizio Bevilacqua
Dipartimento di Ingegneria Industriale e Scienze Matematiche
Università Politecnica Delle Marche
Via Brecce Bianche, 12, 60131, Ancona, Italy
phone: +39 0712204771
e-mail: m.bevilacqua@univpm.it

Received: 21 January 2020
Accepted: 27 July 2020

Abstract
The application of churn prevention represents an important step for mobile communication companies aiming at increasing customer loyalty. In a machine learning perspective, Customer Value Management departments require automated methods and processes to create marketing campaigns able to identify the most appropriate churn prevention approach. Moving towards a big data-driven environment, a deeper understanding of data provided by churn processes and client operations is needed. In this context, a procedure aiming at reducing the number of churners by planning a customized marketing campaign is deployed through a data-driven approach. Decision Tree methodology is applied to drow up a list of clients with churn propensity: in this way, customer analysis is detailed, as well as the development of a marketing campaign, integrating the *individual churn* model with *viral churn* perspective. The first step of the proposed procedure requires the evaluation of churn probability for each customer, based on the influence of his social links. Then, the customer profiling is performed considering (a) individual variables, (b) variables describing customer-company interactions, (c) external variables. The main contribution of this work is the development of a versatile procedure for viral churn prevention, applying Decision Tree techniques in the telecommunication sector, and integrating a direct campaign from the Customer Value Management marketing department to each customer with significant churn risk. A case study of a mobile communication company is also presented to explain the proposed procedure, as well as to analyze its real performance and results.

Keywords
Big Data Analytics, Machine Learning, Probability Estimation Trees, Customer Value Management, ICT sector.

Introduction

In the Industry 4.0 era, with the introduction of IoT (Internet Of Things) and Big Data concepts, the ICT (Information and Communications Technology) sector offers innovative products and services. This is possible because of the increasing availability of customer data that has also created a new fruitful research area within marketing science: Customer Value Management (CVM) [1]. CVM approach focuses on customer big-data analysis to improve the efficiency of marketing campaigns, customizing the actions. Nowadays, due to the increasing competitiveness and constant price-war, reducing the attractiveness and profitability of the ICT sector, a revision of CVM marketing strategies is needed. In the Telecommunication sector, three main strategies have been proposed to generate more revenues [2]: (a) acquire new customers, (b) upsell the existing ones, and (c) increase retention rate (customer loyalty indicator). A common expression in the telecommunication industry is that it costs five times more to acquire a new customer than it does to retain an existing one [3]. Considering the definitions above and the problem addressed in the current paper, Churn prevention, i.e., the implementation of a methodology to prevent customer abandonment, represents a focal point: a company needs to reduce Churn Rate, as well as increase Customer Lifetime Value (CLV), or customer relationship. CLV is an indicator that measures predictable profits based on the relationship with customers; customer Churn Rate indicates customer satisfaction and service quality, it refers to customers losses who switch from one company to another competitor within a given period [4]. Since churn prediction helps the company to do something and retain their customer [5], it should be known in advance [6]. Moreover, customer loyalty can generate positive word-of-mouth, bringing significant advantages over other types of promotion in terms of credibility [7]. Companies are eagerly seeking big data analytics solutions to solve customer churn problems or other aspects such as predictive maintenance [8], for the sake of turning the data into valuable business insights [9].

In this paper, we propose a data analytics model integrating the data collected from different sources, to improve marketing CVM decision-making; a similar perspective has already been implemented in the energy sector [10].

The predictive model developed in this work uses machine learning (ML) techniques and builds a new way of feature engineering, going from the identification of customers churn risk probability to the implementation of marketing campaigns aimed at retention: a procedure to identify the right customer, at the right time, with the right offer, is developed.

In addition, interactions among individuals are studied using Social Network Analysis (SNA) and Classification Tree approaches. According to this theory, customer churn is not only influenced by *endogenous factors* (e.g., the telephone subscription type) but also by *exogenous factors* (e.g., the influence of other individuals). Qualitatively, a client is more likely to "churn" if people having a strong influence on him have migrated to another operator. This is what the authors have called the "**viral effect**", and it must necessarily be considered form the mathematical point of view in a churn prevention perspective. The remainder of the paper is organized as follows: after this brief introduction and a review of the most relevant contributions on similar topics (Sec. 2), the procedure is developed in Sec. 3, additionally describing the processes to plan a marketing campaign. Moreover, in Sec. 4, the case study of a well-known Italian Telecommunication company is presented to clarify better the application of the procedure and the effectiveness of the developed model. Lastly, in the same section, the conclusions and future research directions are presented.

Literature review

Churn Prevention and management models are widely used in CVM departments for customer churn prediction. In the last few years, the use of industrial big data for customer churn management has caught researchers studies because traditional methods are not engineered for the type of big, dynamic, and unstructured data [9]. First of all, [7] provided an understanding of the interplay between switching costs and loyalty. The problem of churn prediction has been tackled in many different sectors, such as telecommunication [2, 5, 11–13], e.g., internet service provider [14], retailing [15, 16], financial industry, e.g., banking [17, 18], energy [19].

Reference [20] identified three different key factors of the customer churn problem in existing literature, namely, (a) churn rate, (b) prediction performance, and (c) retention capability. The customer churn problem can be addressed through different techniques. Several authors, like [21, 23], presented a comparative study of the most used algorithms categories for customer churn prediction: (a) Regression analysis, (b) Decision tree, (c) Bayes algorithm, (d) Support Vector Machine, (e) Instance-based learning (f), Ensemble learning, (g) Artificial neural network, (h) Linear Discriminant Analysis. Reference [2] developed a call-behaviour-based churn-prediction technique from contractual information of the subscribers and call pattern changes. Amin et al. [13] identified that there is still a lack of efficient customer churn prediction approaches in the telecommunications sector; it is estimated that one in fifty subscribers of a given company discontinues their services every month [4]. Table 1 lists several studies relating to the building of Churn Prediction models by using a telecoms dataset. The table shows the authors, the selected modelling techniques, and, finally, their results.

The literature review highlights different studies in the development of Individual Churn model and SNA: the latter technique is applied in several studies for data transformation and preparation, e.g., for environmental risk management [27].

In the ICT sector, SNA is performed to summarize the connections between every two customers and build a social network graph [22] with different types of weights. Part of this, for example, is the effect of "word-of-mouth".

Table 1
Churn prevention techniques in the telecommunications sector.

Ref.	Techniques	Results
[6]	Apache Hadoop, Apache Hbase, C4.5 Algorithm	Used customer call records as input for churn prediction.
[9]	Hadoop MapReduce, Clustering method	Proposed a semantic algorithm stronger than the subtractive clustering method (SCM) and fuzzy c-means (FCM).
[11]	Apache Spark, Hadoop, SQL.	Predicted churn based on the 3V's perspectives of big data: Variety plays the most important role.
[12]	Random Forest	Real-time capture of customer experience to assess which conditions lead the user to call the telco's customer care centre.
[13]	Rough set theory	Four rule-generation mechanisms: Genetic Algorithm is the most efficient technique.
[14]	AdaBoost, Extra Trees, KNN, XGBoost	SMOTE technique before comparing several models: XGBoost gave the best result.
[20]	Receiver Operating Characteristic curve, customer Lifetime Value	Developed a Profit model for churn prevention
[22]	Decision Tree, Random Forest, Gradient Boosted Machine Tree, Extreme Gradient Boosting	Testing and training of four different models without any proactive action: XGBOOST gave the best result.
[24]	Survival analysis techniques	Predicted the membership duration of customers
[25]	Artificial neural networks (ANN), Self-Organizing Maps (SOM)	The hybrid models (ANN+ANN and ANN+SOM) outperform the single neural network baseline model in terms of accuracy
[26]	Decision Tree, Neural Network	Achieved good prediction accuracy using customer profile, billing and call information and service change records

According to existing literature contributions, the research gap involves the development of a **Viral Churn model** with a direct marketing campaign. Due to the wide amount of data nowadays available, the identified research gap could be addressed capitalizing on ML techniques to determine the conditioned churn probability between different customers.

Research approach

The current research approach is a modification of different methods, aiming at developing a procedure to organize items in the CVM department: SNA and Decision-Tree application for "Viral churn" prevention. Specifically, the main steps of a marketing campaign are defined based on a customer loyalty strategy, such as the "actuator" of the churn prevention strategy. The operational process consists of a cyclical multi-layer model (Fig. 2): 1) Data management, 2) decision-making model, 3) marketing campaign application, and 4) performance analytics.

Data management

Data management is the first fundamental step in CVM to define customer profiling as input for the next phase. According to the proposed process, the target is dynamically modified based on the analysis layer improving strategy effectiveness. From these premises, data collection for customer analysis is summarised below. As highlighted in Fig. 1, individual parameters and customer-company interaction are limited only to the intrinsic characteristics of the client, constituting the so-called Individual churn model. The activation and deactivation number are used because users often changing services are more conditionable. Data allowance and usage are included since customers with high usage rate often look for the best offer; traffic and spending control are checked because those who usually monitor parameters such as expenditure, is more sensitive to an offer.

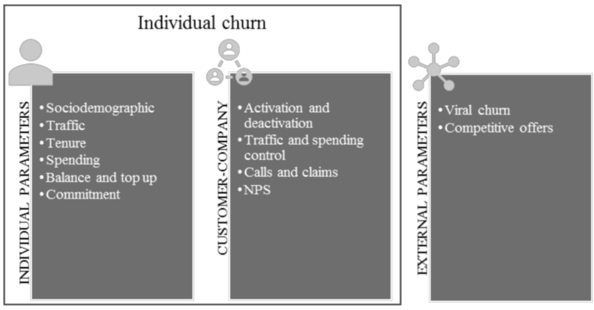

Fig. 1. Customer analytics parameter

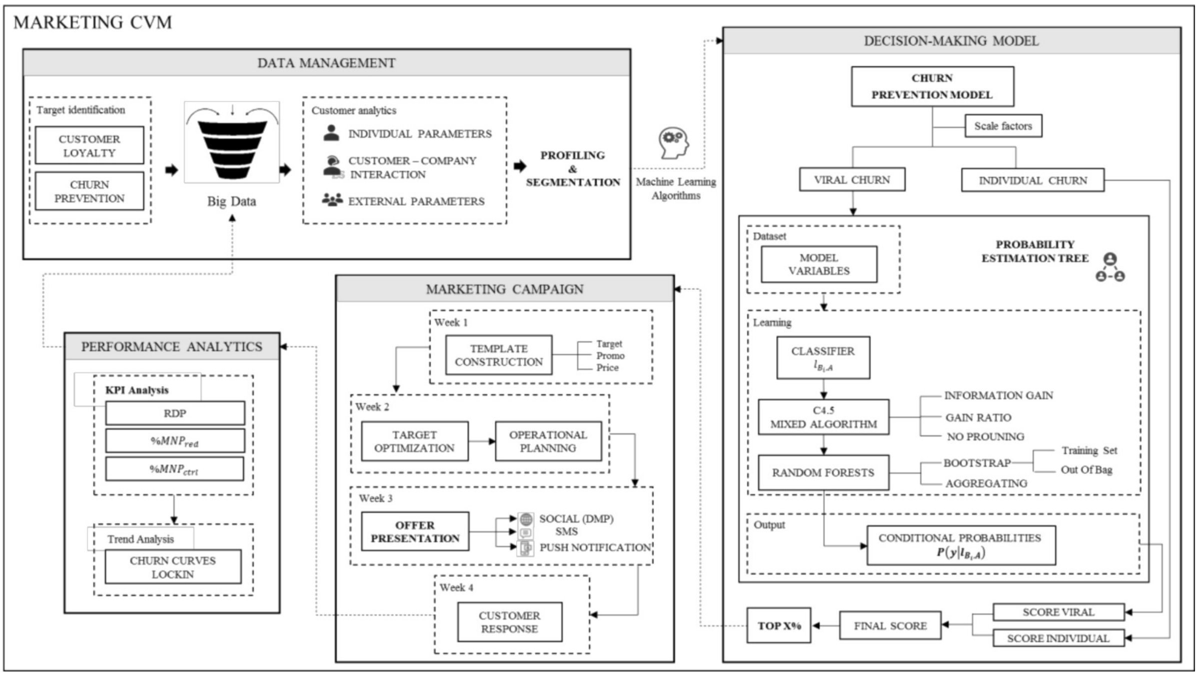

Fig. 2. Marketing CVM framework.

Finally, tenure represents the period of service subscription, while NPS (Net Promoter Score) the management tool used to assess loyalty in a business-to-customer relationship. All these variables are used for customers churn prevention; however, a client could churn simply because people around him have done the same.

In order to improve the analysis executed by previous models, which for ease of implementation and results are still the most widely used, a **Viral Churn model** is developed. In this sense, the *external parameters* concerning competitive offers are used to understand market evolution (e.g., in the telecommunication sector the focus is shifted from increasing company-customer relationship to customer loyalty). At the same time, viral churn variables aim at discovering the correlations among different customers' churn (as in the case of failure analysis [28]).

Decision-making model

The second fundamental step consists of a decision-making model development: the input target is translated into an ML algorithm in order to identify customers at churn risk. Specifically, the output of the model is a daily list with the **top x%** customers potentially churning. Our model consists of three "sub-models": the (a) Individual churn, based on the intrinsic variable of the customers, the (b) Viral churn, based on the SNA in order to measure the so-called "viral" effect (i.e., churners influence), and the (c) Combination of both models, based on scale factors. Since the aim of the work is deepening the intrinsic mechanisms of the Viral churn model, the other ones are treated briefly only to understand their global functioning.

Our Viral churn model consists of three main phases, namely, Dataset, Learning, and Output.

Dataset

The viral churn propensity is based on social networks influence: assuming g customers, g links, and r relationship, the model can be mathematically formalized, as follow:

- $G(N, L, X)$, graph-oriented and weighted;
- N^g vector containing the customer base;
- L^{fxr} vector of all the existing links between the customer pairs, where $l_{ij} = (n_i, n_j)$ indicates the link between node i and node $j (l_{ij} \neq l_{ji})$;
- X^{fxr} matrix containing the values of the model r variables for each link, in which x_{kr} is the value of the r-th relationship type of the pair (n_i, n_j).
- The main variables of the Viral Churn model are listed in Table 2.
- "Churn_y" is the model label (e.g., if $Churn\_y = 0$, probably the churn of customer j is not conditioned by customer i). In our model dataset, only active links are considered.
- Table 2 is composed of symmetrical and asymmetrical variables. The latter expresses the "force" of node i towards node j, decreasing as the time since the churning event increases.

Table 2
Variables of the "Viral churn model".

Name	Description
N_calls	Number of calls between nodes i and j
N_messages	Number of messages between nodes i and j
A_calls	Average call duration between nodes i and j
Position	Average distance between nodes i and j
Churn_y	Dichotomous variable (= 1 if the operator of node i is changed)
Delay	Time elapsed since migration of node i
Operator	Operator to whom the customer has migrated
Degree_centrality	Number of persons with whom node i is directly connected
Betweennes_centrality	Number of minimum paths between two nodes (including node i) divided by the total number of minimum paths between two nodes.

Learning

In Fig. 3, a simplified graph is shown in order to illustrate the intrinsic mechanism of our proposed model, by assuming A as the customer whose churn probability must be calculated, B_i the i-th customer who could influence A having already changed operator, and $l_{B_i,A}$ the link measured by the variable *churn_y*. The conditional probability of churn event A, given churn event B_1 and B_2, is denoted by $P(A|B_1, B_2)$. Assuming n churn events, the churn probability estimate is as follows:

$$P(A|B_1, B_2, ..., B_i, ..., B_n). \tag{1}$$

Since:

- the application of Bayes Theorem,
- the assumption that churn event of customer A is influenced by the "OR" probability of n events,
- the application of the fundamental probability axiom of a finite union of not mutually exclusive events,
- and the application of Poincaré's formula, for computational efficiency and simplicity, the class probability estimation is approximated as:

$$P\left(\bigcup_{i=1}^{n} A|B_I\right) \approx \sum_{i=1}^{n} P\left(A|B_i\right). \tag{2}$$

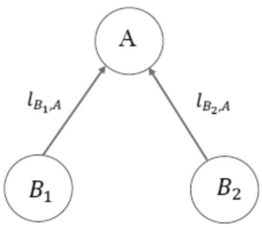

Fig. 3. Viral churn model: simplified graph.

Finally, the viral churn probability is as in (3):

$$P\left(A|B_1, B_2, ..., B_i, ..., B_n\right) = \sum_{i=1}^{n} P\left(A|B_i\right). \tag{3}$$

The conditional churn probability of customer A is closely related to the number of clients "*connected*" B_i and to the "*strength*" $P\left(A|B_i\right)$ of the bond $l_{B_i,A}$. Referring to (3), an estimation of each probability $P\left(A|B_i\right)$ is needed by using ML methods for pattern recognition (Fig. 4): Decision Tree, Probability Estimation Tree, C4,5 Algorithm revised, and Random forests.

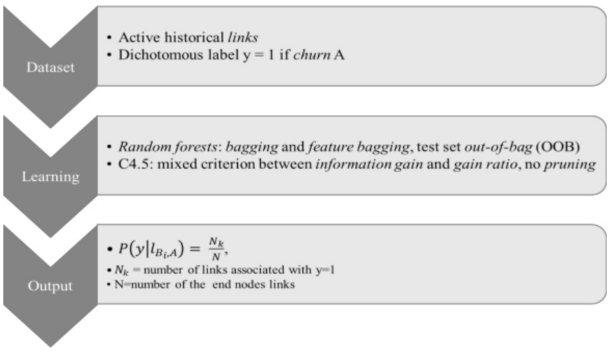

Fig. 4. Phases of the viral churn model.

A **Decision Tree (DT)** is a *classification* model learning from a *training set* $\{l_{B_i,A}, y\}$, where $l_{B_i,A}$ is the attributes vector, and $y \in \{0, 1\}$ is the *dichotomous class-label* vector. The unlabeled sample is categorized into class $y = 1$ if it falls into the decision areas corresponding to $y = 1$. The implemented DT is a decision support tool that uses a tree-like graph in which the internal nodes represent a "test" on an attribute $l_{B_i,A}$. Each branch represents the outcome of the test, and each end-node represents a class label (classification taken after computing all attributes). The classification rules are represented by the paths from the root to leaf (end-node). Finally, the accuracy of our predictive model is verified through a *test set* from the comparison between the real class and the assigned class. The role of our classifier is both assigning a class to the input links and predicting the conditional class probability $P(y = 11)$ for each customer. DT models do not usually estimate the Class Probability, but, in our model, big data are analyzed to identify both the association and the conditioned probability between linked customers.

In this sense, the **Probability Estimation Tree (PET)** is chosen as a decision boundary-based theory for assigning probabilities of each unlabeled sample that falls into a leaf. The belonging probability of each class is estimated considering the highest likelihood; since the homogeneous classes of the Algorithm, more end-nodes than classes are created bygrouping the links according to their similarity. Equation (4) gives a formal expression:

$$P\left(y=1|l_{B_i,A}\right)=\frac{N_k}{N}, \tag{4}$$

N_k is the number of samples from the class k (i.e., the number of links associated with A's churning), compared to the total number of data in the leaf.

Three main aspects are handled: (i) the estimates are simplistically shifted towards zero or one considering that in our case-study the average monthly churn is about 1–2%; (ii) the training set is often inaccurate if divided into small subsets, conversely, our data set is around 600,000 units; (iii) the accuracy of our classification model is improved by the **C4.5 revised Algorithm**: a greedy top-down inductive learning approach.

Step 1. The initial node of the tree consists of both the training data set (S) and the attributes set or variables.

Step 2. Stop conditions evaluation in order to identify the tree lefts (or end-nodes):
- All the elements with the same label.
- Empty node (parent node label assignment).
- Using all splitting attributes (most frequent class label assignment).

Step 3. Selection of the partitioning attribute with the greatest reduction in entropy and purest nodes production, measured by the "entropic gain". Since the measure of Information gain selects the attributes with a large number of distinct values, the sub-trees could be repeated several times in a single decision tree reducing the classification efficiency. As a result, in our Algorithm a mixed criterion is used to measure the entropic gain (5) by integrating both *information gain* and *gain ratio*: the only attributes with a gain higher than the average value are considered to select the attribute with the highest gain ratio

$$\begin{aligned}\mathrm{Gain}\left(S,A\right)&=\mathrm{Entropy}(S)\\&-\sum_{v\in \mathrm{Value}(A)}\frac{|S_{A=v}|}{|S|}\mathrm{Entropy}\left(S_{A=v}\right),\end{aligned} \tag{5}$$

$$\mathrm{Entropy}(S)=-p_{-}\log_2\left(p_{-}\right)-p_{+}\log_2\left(p_{+}\right). \tag{6}$$

Entropy(S) measures the quantity of uncertainty of the whole S, where p_+ is the proportion of the class examples $p=1$ (i.e., the links to customers who have churn) contrary to p_-. Entropy$(S_{A=v})$ is instead the conditioned entropy, calculated in the subset of all the examples with the attribute A of value v. Since it is not possible to evaluate all the continuous attributes values (as opposed to discrete attributes), the examples i are sequenced according to the increasing value of the attribute to be evaluated, then, the maximum gain ratio corresponding to the "threshold" c (i.e., the average between two consecutive values) is compared with the other attributes gain. The attribute A with maximum gain is selected for partitioning

$$c_i=\frac{x_i+x_{i+1}}{2}\quad \forall i\in A. \tag{7}$$

Step 4. The node A is partitioned into subsets corresponding to all its values (if discrete), or in two sub-sets ($A <= c$; $A > c$), with "cut-off point" just the value c corresponding to the maximum gain (if A is continuous).

Step 5. The Algorithm is applied iteratively for each node identified, excluding from the analysis the attribute or the threshold c, until one of the stop conditions in Step 2 is satisfied.

In our model, pre-pruning techniques are excluded because, while increasing the algorithm efficiency by reducing overfitting, no optimal solutions can be found contrasting with the PET target. Post-prouning techniques are also excluded aiming to "cluster", as many classes as possible, in order to obtain more accurate estimates. Aiming to prevent the problem of a decision-tree with reduced bias and high variability, **Random forests** are developed based on two techniques: *Bagging* and *Feature Bagging*. Bagging consists of the "Bootstrap" method and the "aggregating" alghoritm, for resampling the dataset by extracting m samples of quantity n as training-test, and the other ones as test-set (out-of-bag estimate, OOB). Multiple different models from a single training dataset are created, reducing the variability for the Central Limit Theorem. Based on Feature Bagging, the variability of a random forest decreases as the number of trees in it increases. In fact, if one or more attributes are very strong predictors for classification are selected in several trees, causing the correlation of the outputs and the subsequent reduction of model variability.

Output

Once the decision trees are developed, the estimate of the final probability (8) is the average of the conditional probabilities of the individual decision tree

$$P\left(y|l_{B_i,A}\right)=\frac{1}{m}\sum_{j=1}^{m}P_j\left(y|l_{B_i,A}\right). \tag{8}$$

The weighed combination between Viral Churn and Individual Churn value is provided by the final score (Fig. 5) for each customer, as the summation of all the probabilities of the associated links (9)

$$\text{Score}=\sum_{i=1}^{n}P\left(y|l_{B_iA}\right). \tag{9}$$

The final output of this phase is a **list of the top x%** of customers at churn risk.

Fig. 5. Customer score construction

Marketing campaign

CVM philosophy is applied through standardized and automated business processes, aiming at the creation of a personalized campaign for customer loyalty. The main features of our model are:

- monthly planning,
- two-week lead time for lists generation,
- human resources as bottleneck,
- weekly report about the progress of the campaigns.

The creation process of our marketing campaign consists of four weeks, as summarized below.

Week 1. Creation of the initial template and target definition. The template consists of general information about the campaign, such as name and identification code, purpose, duration, service, communication channel (sms, push notification, or social/web), activation channel, text, and offer description. The target lists from the Viral churn model are refined by an optimizer, according to policy regulations or qualitative aspects (translated in Boolean language).

Week 2. Creation of the final template and operational planning (submissions are distributed within five working days, considering any system capacity constraints).

Week 3. Presentation of the offer.

Week 4. Customer response. The uniqueness of the campaign identification code combined with the target number (from 1 to 8) generates the so-called *keycode* to identify the offers activated.

Once the campaign is officially over, all data are extracted and collected automatically from the information systems of company databases. This is the input for the analysis phase, namely the feedback performance of the marketing campaign.

Performance analytics

As in the case of a new product projects development [29], performance analysis is essential to study the benefits of each marketing campaign through Key Performance Indicators (KPI). In terms of churn prevention, the campaign efficiency is measured by the number of loyal customers through the service offered. In this sense, our first KPI is the *trade-off* between the campaign efficiency and the incurred expenses.

In order to measure the campaign efficiency, the following data are analyzed: *campaign numerousness*, *redeemers*, i.e., customers who activate the service, *mobile number portability* (MNP OUT), i.e., customers who change operator.

Campaign efficiency is measured by the percentage *RDP*, directly proportional to the effectiveness of profiling, offering, and communication

$$RDP=\frac{\text{redeemers}}{\text{CustomerNumberInCampaign}}. \tag{10}$$

Communication usefulness is measured through the *Churn reduction* by the Scenario Analysis comparing the redemption (customers lost) of the scenario DO (campaign test) and the scenario DO NOTHING (a sample of 5% of the customers of the list are exluded in order to test the campaign)

$$\text{churn reduction}=\frac{\%\,\text{MNP}_{\text{red}}*\text{redeemers}-\%\text{MNP}_{\text{ctrl}}*\text{redeemers}}{\%\text{MNP}_{\text{ctrl}}*\text{redeemers}}, \tag{11}$$

$\%\,\text{MNP}_{\text{red}}$ is the percentage of redeemers in the test who changed operators, while $\%\,\text{MNP}_{\text{ctrl}}$ is the percentage of customers in the control who have changed operators. These parameters are inversely proportional to the effectiveness of the offer, in fact, an action is effective if:

- $\%\,\text{MNP}_{\text{ctrl}} > \%\,\text{MNP}_{\text{red}}$
- and *churn reduction* <0.

Case study

In this section, the case study of a well-known Italian Telecommunication company is analyzed based on the following assumptions:

- Mobile costumer base.
- Data-set with all the ex-customers who have changed operator in the last three months, assuming that the influence of a customers decreases as the elapsed time from the churn action increases.
- Monthly update of both dataset and MNP graph, in order to adapte better our model with customers changes.
- Daily update of variable values in order to respect the CVM marketing approach of a continuous targetization, maximizing the effectiveness and efficiency of each campaigns.
- Monthly structural change of variables, removing corrupted, incomplete or extreme value data in order to ensure data set integrity.
- Churn lists creation depending on both Individual and Viral churn model. The individual score is assigned to the entire customer base, while the viral score only to customers with active links (i.e., connected to the churners of the previous three months).
- Churners and churn rate update within one month from the list generation, in order to define churn propensity in a limited time.

Results

Churn propensity is influenced by external variables as emphasized below. In Fig. 6 the effectiveness of the targetization model is shown: groups of customers with up to 3 times the average churn propensity are identified. Within the list of churn risk profiles, approximately 2.8 million customers with the double average risk of the entire customer base are identified, as a starting subset for the identification of churn prevention campaigns targets. Of the 600 thousand affected by virality, 120 thousands (about 20% of the customers identified) are within the top 5%, i.e., high risk customers, equal to about 2.8 million of the global customer base. Once the customers are identified, a subset of 12,000 customers of the 120,000 having high-risk of churn are taken to start a test campaign offering free Giga for 3 or 6 months, depending on the target.

Risk Rank	Customers	Churners	Churn rate
Top 5%	664,358	20,670	*3.11%*
Top 10%	703,262	17,593	*2.50%*
Top 15%	725,785	15,893	*2.19%*
Top 20%	739,141	14,693	*1.99%*
CUSTOMER BASE	15,065,853	156,171	*1.04%*

Fig. 6. List profiling performance table.

After one week of campaign, the redemption is 19% and MNP −77%. Considering the average RDP between 2–6% in a standard campaign, our result is optimal and the negative value of number portability confirms the effectiveness of the marketing action for customer loyalty. Based on the excellent performance, our marketing campaign is applyed weekly for 5 mounths to the entire churn risk customer base, obtaining an average RDP 14%. In Figs 7 and 8, the average monthly results of $\%MNP$ and churn reduction are shown respectively.

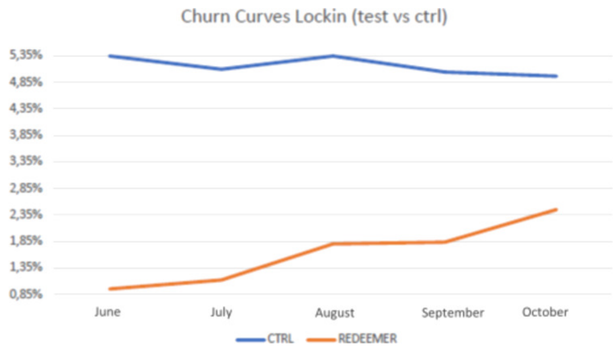

Fig. 7. Trend of %MNP (control and test).

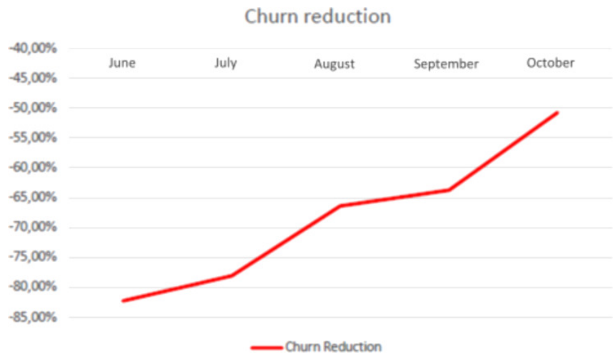

Fig. 8. Trend of churn reduction.

As shown in Fig. 7 and Fig. 8, after the optimal effect of initial churn prevention, the values decreased rapidly, still achieving an extremely high performance campaign. The "positive" trend of the churn reduction suggests that the effect could stop in a few months, however, as long as the parameters are "good" ($RDP \sim 5\%$, $Churn\ reduction > 0$) the campaign will remain in production. To prevent the performance decrease, continuous model improvement it is necessary for identifying potential churners accurately.

Discussion

In this section, the main points of our paper and the possible future solutions are discussed.

Decision Tree is used to make decisions even if sometimes it is applied for class probabilities estimation. Usually, DT model is based only on internal factors (e.g., calls number and usage) or SNA (e.g., investigating social structure). Instead, in this paper, DT is applied to understand how one customer churn influences the choices of another by providing excellent results.

Specifically, random forests are applied in our model, combined with C4.5 Algorithm. These two approaches give low bias (C4.5) as well as a low variability (random forests approach), providing a difficult interpretation model. Several variables are identified for daily customer targetization, providing efficient churn risk lists to better analyzed an extremely competitive market. The conditioned probability, based on the union of churn events, declines statistical relevance.

However, the aim of the paper is not to have a precise and accurate churn probability estimation, but to identify measurable ranking customers based on the *trade-off* between output effectiveness and computational efficiency.

In order to suggest future applications and extensions for continuous improvement, different solutions are proposed:

a) model structure update to improve the effectiveness of profiling;
b) new commercial offers differentiated according to the target group;
c) automatic daily campaign creation by integrating the extraction process and target creation. Target ranking could be optimize, providing directly customized offers through daily schedule, minimizing human errors and their associated costs.

Conclusions

The telecommunication marketing aims to exploit the data collected on customers (average spending, purchasing habits, use of services, etc.) to improve business relations by focusing on customer loyalty. Continuous improvement is the basis of CVM marketing philosophy, involving not only model development but also marketing campaign processes. This paper is directed to churn prevention by combining both Individual and Viral churn model to identify the conditioned probability that a customer will make churn influenced by a linked customer (who has already made churn). Our study consists of customer profiling, target segmentation and results analysis of a real ICT company by using DT models and personalized marketing campaign.

The purpose of the paper is to maximize the effectiveness (number of service activations) and efficiency (process completion time and minimization of human errors) of the marketing actions. The readiness of marketing campaigns improves as process efficiency increases, leading to competitive advantage in terms of responsiveness to market variability.

In conclusion, the Viral effect does not only depend on the type of link but also on the number of active links: the more people close to a customer change operators, the more likely she/he will churn.

References

[1] Verhoef P.C., Doorn J., Dorotic M., *Customer value Management: An overview and research agenda*, Marketing ZFP, 29, 2, 105–122, 2015.

[2] Wei C.P., Chiu I.T., *Turning telecommunications call details to churn prediction: a data mining approach*, Expert Syst. Appl., 23, 2, 103–12, 2002.

[3] Khan M.R., Manoj J., Singh A., Blumenstock J., *Behavioral modeling for churn prediction: Early indicators and accurate predictors of custom defection and loyalty*, IEEE International Congress on BigData, pp. 677–680, 2015.

[4] Lu N., Lin H., Lu J., Zhang G., *A customer churn prediction model in telecom industry using boosting*, IEEE Trans. Ind. Informat., 10, 2, 1659–1665, 2014.

[5] Tiwari A., Sam R., Shaikh S., *Analysis and prediction of churn customers for telecommunication industry*, International conference on I-SMAC, pp. 218–222, 2017.

[6] Limaye G.D., Chaudhary J.P., Punjabi S.K., *Churn Prediction using MapReduce and HBase*, International Journal on Recent and Innovation Trends in Computing and Communication, 3, 3, 1699–1703, 2015.

[7] Caruana A., *The impact of switching costs on customer loyalty: A study among corporate customers of mobile telephony*, Journal of Targeting, Measurement and Analysis for Marketing, 12, 3, 256–268, 2003.

[8] Antomarioni S., Bevilacqua M., Potena D., Diamantini C., *Defining a data-driven maintenance policy: an application to an oil refinery plant*, International Journal of Quality & Reliability Management, 36, 3, 2018.

[9] Bi W., Cai M., Liu M., Li G., *A Big Data Clustering Algorithm for Mitigating the Risk of Customer Churn*, IEEE Transactions on Industrial Informatics, 12, 3, 1270–1281, 2006.

[10] Bevilacqua M., Ciarapica F.E., Diamantini C., Potena D., *Big data analytics methodologies applied*

at energy management in industrial sector: A case study, International Journal of RF Technologies: Research and Applications, 8, 3, 105–122, 2017.

[11] Huang Y., Zhu F., Yuan M., Deng K., Li Y., Ni B., Dai W., Yang Q., Zeng J., *Telco Churn Prediction with Big Data*, Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data, pp. 607–618, 2015.

[12] Diaz-Aviles E., Pinelli F., Lynch K., Nabi Z., Gkoufas Y., Bouillet E., Calabrese F., *Towards real-time customer experience prediction for telecommunication operator*, International Conference on Big Data (Big Data), pp. 1063–1072, 2015.

[13] Amin A., Anwar S., Adnan A., Nawaz M., Alawfi K., Hussain A., Huang K., *Customer churn prediction in the telecommunication sector using a rough set approach*, Neurocomputing, 237, 242–254, 2016.

[14] Do D., Huynh P., Vo P., Vu T., *Customer Churn Prediction in an Internet Service Provider*, IEEE International Conference on Big Data, pp. 3928–3933, 2017.

[15] Subramanya K.B., Somani A., *Enhanced Feature Mining and Classifier Models to Predict Customer Churn for an E-retailer*, Graduate Theses and Dissertations, 2016.

[16] Shirazia F., Mohammadib M., *A big data analytics model for customer churn prediction in the retiree segment*, International Journal of Information Management, 48, 238–253, 2019.

[17] Cui S., Ding N., *Customer Churn Prediction Using Improved FCM Algorithm*, 3rd International Conference on Information Management, pp. 112–117, 2017.

[18] Sayed H., Abdel-Fattah M.A., Kholief S., *Predicting Potential Banking Customer Churn using Apache Spark ML and MLlib Packages: A Comparative Study*, International Journal of Advanced Computer Science and Applications, 9, 11, 674–677, 2018.

[19] Moeyersoms J., Martens D., *Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector*, Decision Support Systems, 72, 72–81, 2015.

[20] Zhang Z., Wang R., Zheng W., Lan S., Liang D., Jin H., *Profit Maximization Analysis Based on Data Mining and the Exponential Retention Model Assumption with respect to Customer Churn Problems*, IEEE 15th International Conference on Data Mining Workshops, pp. 1093–1097, 2015.

[21] Sabbeh S.F., *Machine-Learning Techniques for Customer Retention: A Comparative Study*, International Journal of Advanced Computer Science and Applications (IJACSA), 9, 2, 273–281, 2018.

[22] Ahmad A.K., Jafar A., Aljoumaa K., *Customer churn prediction in telecom using machine learning in big data platform*, Journal of Big Data, 6, 28, 2019.

[23] Mahajan V., Misra R., Mahajan R., *Review of Data Mining Techniques for Churn Prediction in Telecom*, JIOS 2015, 37, 2, 183–197, 2015.

[24] Junxiang L., *Predicting Customer Churn in the Telecommunications Industry – An Application of Survival Analysis Modeling Using SAS*, Data Mining Techniques, 2002.

[25] Tsai C.F., Lu Y.H., *Customer churn prediction by hybrid neural networks*, Expert Syst. Appl., 36, 12547–12553, 2009.

[26] Hung S.Y., Yen D.C., Wang H., *Applying data mining to telecom churn management*, Expert Syst. Appl., 31, 515–524, 2006.

[27] Ciarapica F.E., Bevilacqua M., Antomarioni S., *An approach based on association rules and social network analysis for managing environmental risk: A case study from a process industry*, Process Safety and Environ. Protection, 128, 50–64, 2019.

[28] Antomarioni S., Pisacane O., Potena D., Bevilacqua M., Ciarapica F.E., Diamantini C., *A predictive association rule-based maintenance policy to minimize the probability of breakages: application to an oil refinery*, International Journal of Advanced Manuf. Technology, pp. 1–15, 2019.

[29] Ciarapica F.E., Bevilacqua M., Mazzuto G., *Performance analysis of new product development projects: An approach based on value stream mapping*, Int. Journal of Productivity and Perf. Management, 65, 2, 177–206, 2016.